Classification of Sleep-disordered Breathing

JEAN-JACQUES HOSSELET, INDU AYAPPA, ROBERT G. NORMAN, ANA C. KRIEGER, and DAVID M. RAPOPORT

Division of Pulmonary and Critical Care Medicine, New York University School of Medicine, New York, New York

Increasing recognition of sleep-disordered breathing (SDB) and its morbidity have prompted reevaluation of techniques to identify respiratory events during sleep. The present study was designed to evaluate the utility of various metrics of SDB and to identify the optimal respiratory metric that objectively correlates to symptoms of excessive daytime somnolence (EDS). Metrics were derived from combinations of conventional apnea/hypopnea, flow limitation events (transient elevated upper airway resistance identified by characteristic flattening on the flow/time tracing, using a noninvasive nasal cannula technique), desaturation, and arousal. A total of 137 subjects underwent clinical evaluation and nocturnal polysomnogram. In 34 randomly selected subjects, the best metrics for discriminating between 13 subjects with no EDS/snoring and 21 patients with EDS and snoring were identified by receiver operator curve analysis. Of the metrics and cut points tested, a total respiratory disturbance index (RDI_{Total}, sum of apneas, hypopnea, and flow limitation events) of 18 events/h was found to have the best discriminant ability (100% sensitivity and 96% specificity). Prospective testing of this metric was then performed with the remaining 103 subjects (14 nonsnoring non-EDS, 21 snoring non-EDS, 68 snoring with EDS). Using this cutoff of 18 events/h, we obtained 71% sensitivity and 60% specificity for identifying subjects with EDS. We conclude that, in subjects with upper airway dysfunction, an index that incorporates all respiratory events provides the best quantitative physiological correlate to EDS.

In the past decade there has been increasing recognition of the high prevalence of obstructive sleep apnea syndrome in both the clinic and general populations (1). It has been proposed that even mild degrees of sleep-disordered breathing (SDB) may be associated with significant morbidity, including excessive daytime somnolence (2–4), long-term cardiovascular complications (5, 6), and significant societal costs (7–9). These public health considerations, as well as increased public awareness of the possible medical significance of SDB, have led to an exponential increase in referrals to sleep disorders centers, especially for evaluation of snoring.

There is still much disagreement on the type and level of respiratory abnormality that must be used as a "cutoff" for significant disease (e.g., obstructive sleep apnea syndrome [OSAS] and upper airway resistance syndrome [UARS]). In a publication by the American Academy of Sleep Medicine, UARS and OSAS were defined as > 5 respiratory events/h and the presence of excessive daytime somnolence (EDS) (10). Part of the problem lies in not knowing the degree of biological variation in the number of respiratory events that exists in subjects who are asymptomatic and have no evidence of long-term morbidity ("normal" subjects). While this problem is beginning to be addressed in large epidemiological populations (11), any attempt to quantify sleep-disordered breathing is complicated by the use of different techniques to detect respiratory events.

Until recently, the standard tool used for monitoring respiratory airflow in polysomnography has been the thermistor. This device is effective at identifying complete cessation of airflow (apnea). However, detection of intermittent reduction of flow (hypopnea) is much more subjective; it requires an arbitrarily chosen reduction in flow and has been defined with and without the need for confirmatory consequences such as desaturation or an arousal (10, 12–14).

Furthermore, the thermistor is insensitive to other sleep-disordered breathing events that may complete the spectrum of upper airway dysfunction associated with sleep (15, 16). Work with more intensive monitoring (esophageal manometry) or using a more sensitive analysis of traditional signals (respiratory inductance plethysmography plus thermistor [17]) has suggested the existence of a large population whose symptoms appear to be due to frequent upper airway events similar in pathophysiology to apnea, but that are undetected by the thermistor (2, 18). The important conceptual advance provided by esophageal manometry has revolutionized the clinical approach to "unexplained" excessive daytime somnolence but remains limited in clinical applicability by its invasiveness and the increased patient discomfort necessary to detect the respiratory abnormality.

In part because of the ambiguity in the definition of respiratory events, treatments to relieve symptoms of snoring and excessive daytime somnolence are being applied on clinical grounds with only indirect proof of any respiratory abnormality. Thus, for example, frequent arousals or response to continuous positive airway pressure (CPAP) are used as evidence that upper airway disease was present and of respiratory origin (19). For a more rational approach to clinical and epidemiological issues, it has become imperative to find a simple reliable diagnostic technique to separate patients from individuals who are free of disease. Finding such a metric requires making the assumption that it is possible, using the metric, to explain the level of symptoms (such as EDS) present in a given individual. Any discrepancy between the value of this metric and the level of symptoms in an individual is then attributed to imperfection of the metric. An alternative to this conceptualization is that even with the most ideal metric for representing the physiology of sleep-disordered breathing, differing sensitivity of individuals to the same level of respiratory stress results in differing levels of symptoms. The problem facing researchers is that identifying the relative merit of these two paradigms cannot be accomplished until the "best" metric has been identified to describe the physiology.

We have previously shown that a simple noninvasive tool exists that can both identify apnea/hypopnea and also detect additional respiratory events associated with flow limitation (15, 16) and thus potentially add to the definition of the physiology. Flow limitation is a temporal pattern of airflow that indicates abnormal upper airway resistance and usually results in arousal (20). We have shown that the nasal cannula system detects the same events classified as respiratory effort-related arousals (RERAs), using esophageal manometry (21). The present study was designed to evaluate the utility of various metrics of SDB derived from combinations of events detected

(*Received in original form August 26, 1998 and in revised form May 22, 2000*)

Supported by grants from NIH-NHLBI HL53931 and NCRR M01 RR00096, Mallinckrodt Nellcor Puritan Bennett, the Foundation for Research in Sleep Disorders, and the American Lung Association of New York.

Correspondence and requests for reprints should be addressed to David M. Rapoport, M.D., Department of Medicine, New York University Medical Center, 550 First Ave., New York, NY 10016. E-mail: david.rapoport@med.nyu.edu

Am J Respir Crit Care Med Vol 163. pp 398–405, 2001
Internet address: www.atsjournals.org

by the nasal cannula (conventional apnea/hypopnea and flow limitation events), desaturation and arousal. The goal was to determine the optimal metric to objectively correlate symptoms of EDS with their respiratory causes.

METHODS

Subjects

All patients referred to the New York University Sleep Disorders Center (New York, NY) from January 1998 to October 1999 for evaluation of significant EDS or severe snoring (with or without EDS) were considered for inclusion in the present study, supplemented by a group of specifically recruited normal subjects from the community. All patients whose complete sleep history or physical examination strongly suggested a primary nonrespiratory cause of EDS were excluded (symptoms strongly suggesting narcolepsy, poor sleep hygiene, etc.). All patients who presented with predominant symptoms of snoring and did not complain of EDS were included (n = 21). For patients presenting with EDS and thus were clinically suspected to have sleep-disordered breathing, the initial nocturnal polysomnogram (NPSG) was reviewed. We discarded all patients with severe obstructive apnea syndrome; that is, an apnea/hypopnea index (AHI) > 45/h. For this purpose, apnea/hypopnea was defined as a > 50% reduction in airflow lasting > 10 s. These patients were eliminated because the focus of the present study was on the clinically relevant question of determining the physiologic metric most relevant to EDS in UARS and mild/mod OSAS, and we did not wish to bias our statistics with easily identifiable severe OSAS. In total, we retained 89 patients with suspected sleep-disordered breathing. An additional 27 normal subjects were recruited who had no history of either EDS, snoring, or any other sleep complaints. This resulted in a total of 137 subjects whose data were analyzed. The protocol was approved by the New York University Institutional Board of Research Associates.

Evaluation of Symptoms at Baseline

All 110 clinical patients in the present study sought medical attention for symptoms that were clinically attributed to SDB (89 for EDS and snoring and 21 for snoring alone). EDS was defined as a global clinical impression by the sleep clinician of inappropriate and excessive tendency to fall asleep. This was based on patient and/or bedpartner interview at the time of clinical intake and prior to the NPSG. Global clinical impression was felt to best characterize the complaint that causes patients to seek medical attention, and to represent the typical entry criterion for the diagnosis of OSAS/UARS. The Epworth Sleepiness Scale (ESS) (22) was recorded in each patient and Multiple Sleep Latency Test (MSLT) (23) or Maintenance of Wakefulness Test (MWT) was performed as clinically indicated (in 45 patients), but neither ESS nor MSLT/MWT was used to define the presence of EDS. Snoring was defined by a subject's (or bedpartner's) report of loud snoring on most nights of the week.

Development Set

To evaluate potential metrics of SDB and to determine the best cut points for each metric, data from 34 subjects were randomly selected without stratification. This group consisted of 13 nonsnorers without EDS and 21 patients with suspected sleep-disordered breathing (based on the history of severe snoring and EDS). No snorers without EDS were included in the development set.

Test Set

To test the utility of the best metrics developed above (and the associated cut points), we prospectively evaluated the data of the remaining 103 subjects. This group consisted of 14 nonsnorers without EDS, 21 snorers without EDS, and 68 patients with suspected SDB. The "test" set thus contained two groups of subjects who had been clinically classified prior to NPSG: subjects without EDS (nonsnoring or snoring) and patients with EDS and snoring.

Sleep Studies

A full NPSG was recorded in each subject at the New York University Sleep Disorders Center. Recordings of central and occipital electroencephalogram (EEG), electrooculogram (EOG), and submental electromyogram (EMG) were used to monitor sleep. Sleep was scored using the criteria of Rechtschaffen and Kales (24). Leg movements were monitored with an anterior tibialis EMG. A unipolar electrocardiogram (ECG) was used for cardiac monitoring. Oxygen saturation was monitored with a pulse oximeter. Chest wall and abdominal movement were monitored with piezoelectric strain gauges. Respiratory airflow was simultaneously monitored with a nasal/oral thermistor and a

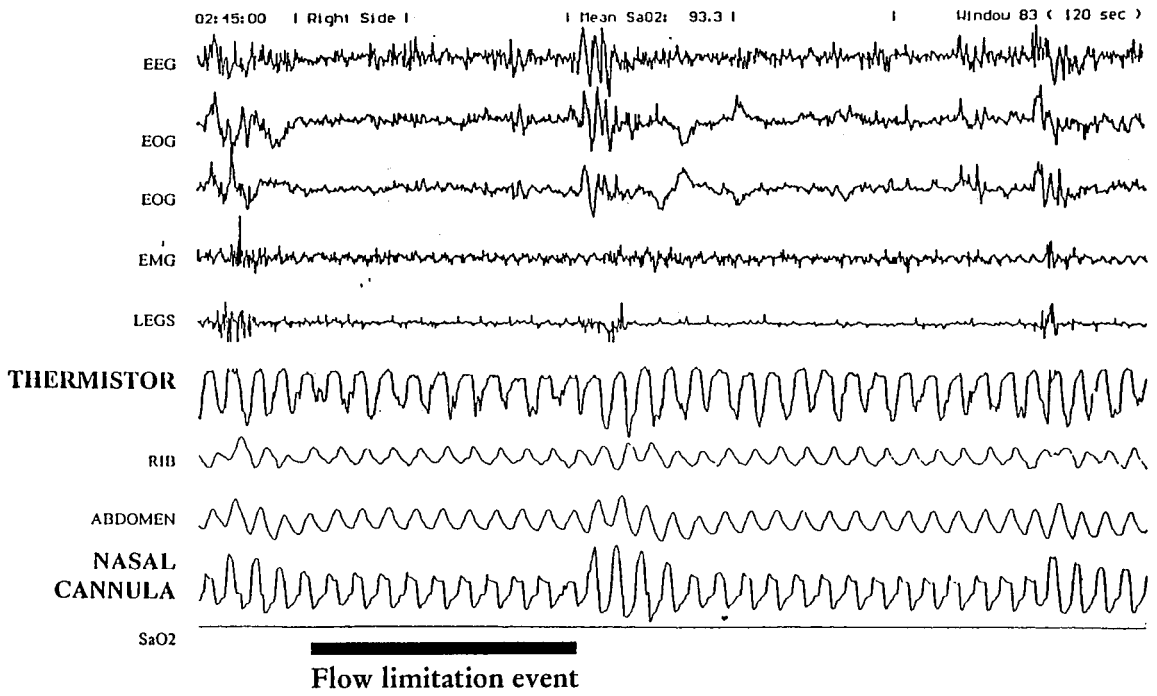

Figure 1. Example of flow limitation event. Note the period of breaths with flattened inspiratory flow contour (marked by the bar) and the rapid return of the flow contour to a sinusoidal shape on arousal from sleep.

TABLE 1

DEFINITION OF SDB INDICES BY THE EVENTS INCLUDED

		Flow Hypopnea		Flow Limitation Events		
	Apnea	With O_2 Desaturation	Without O_2 Desaturation	With O_2 Desaturation	With Arousal	With No Desaturation/Arousal
AHI_{Flow}	x	x	x			
$AHI_{Chicago\text{-}Cannula}$	x	x	x	x	x	
RDI_{Desat}	x	x		x		
$RDI_{Arousal}$	x	x	x		x	
RDI_{Total}	x	x	x	x	x	x

nasal cannula connected to a 2-cm H_2O pressure transducer (Protech [Minneapolis, MN] or Validyne [Northridge, CA]) as previously described (15, 16, 25).

Respiratory events were identified on the basis of nasal cannula flow signal and defined as follows:

Apneas: Absence of airflow on the nasal cannula and < 10% baseline fluctuations on the thermistor signal, lasting for > 10 s

Flow hypopneas (based on airflow measurement only): Airflow amplitude on the nasal cannula < 50% of surrounding baseline, lasting for > 10 s. No O_2 desaturation or arousal confirmation was required to score an event

Flow limitation events: Any series of two or more breaths (lasting > 10 s) that had a flattened or nonsinusoidal appearance on the inspiratory nasal cannula flow signal and ended abruptly with a return to breaths with sinusoidal shape (15, 16). An example of such an "event" is shown in Figure 1. To be counted as a "flow limitation event," the individual breaths had to have a peak flow between 50 and 70% of the peak flow of the surrounding baseline (and thus not be a flow hypopnea)

American Academy of Sleep Medicine (AASM) hypopneas: As proposed by the AASM Task Force (10), these events include both flow hypopneas and any flow limitation event associated with 3% desaturation or associated with an AASM arousal

From these types of events, the following indices of SDB were defined by dividing the relevant number of events by the total sleep time (TST). Table 1 shows which events are included in each index:

AHI_{Flow}: Calculated from the sum of all apneas and all flow hypopneas

$AHI_{Chicago\text{-}Cannula}$: Calculated from the sum of all apneas and all AASM hypopneas. This differs from the AHI_{Flow} in that it includes some additional flow limitation events when these are "upgraded" to a hypopnea by being associated with an arousal or a 3% desaturation

RDI_{Desat}: Calculated from the sum of all apneas and only those nonapnea events (flow hypopneas and flow limitation events) with O_2 desaturation. This differs from the $AHI_{Chicago\text{-}Cannula}$ because it does not include flow hypopneas without desaturation or flow limitation events with arousal only

$RDI_{Arousal}$: Calculated from the sum of apneas, all flow hypopneas, and only flow limitation events with arousal. This differs from the $AHI_{Chicago\text{-}Cannula}$ only in that it does not include flow limitation events with desaturation but no arousal (of which there are few). It is included because in a previous publication (21) we suggested this index

RDI_{Total}: Calculated from the sum of apneas, all flow hypopneas and all flow limitation events without regard to whether there is associated desaturation or arousal

In addition to the above-described analyses based on using a single index to describe SDB, we generated a bidimensional analysis using the apnea index (AI) and the sum of all other events (RDI_{Total}–AI) separately. This was motivated by our previously published preliminary data (26).

Statistical Analysis

A receiver operator curve (ROC) was constructed for each index for the development set, and cut points were chosen to optimize the separation between EDS and no-EDS groups. Sensitivity and specificity along with their 95% confidence intervals were then calculated. The optimal cutoff for separating subjects with and without EDS was then chosen for each index to simultaneously maximize the sum of sensitivity and specificity. The "best" two metrics (maximal area under the ROC) were chosen from among the indices evaluated. The test set was then evaluated prospectively, using only the best indices.

RESULTS

Demographic and polysomnographic data are shown for the development set in Table 2A and for the test set in Table 2B. Both sets of subjects are typical of those seen in clinical practice and span the spectrum from normal to mild–moderate SDB.

Figure 2 shows plots of the six respiratory indices we defined (AI, AHI_{Flow}, RDI_{Desat}, $RDI_{Arousal}$, $RDI_{Chicago\text{-}Cannula}$ and RDI_{Total}). For each index, the left graph shows raw data for both development set and test sets. Three groups (nonsnorers without EDS, snorers without EDS, and snorers with EDS) are indicated. The cutoff (based on the development set alone), which separates those with EDS from those without EDS, is shown by the dashed line. On the right is shown the ROC curve for that index that was used to choose this cut

TABLE 2

DEMOGRAPHIC AND POLYSOMNOGRAPHIC DATA FOR THE DEVELOPMENT SET AND FOR THE TEST SET

Group	No.	Sex (*f/m*)	Age (*yr*)	BMI (kg/m^2)	ESS	TST (*min*)	AI (*per hour*)	AHI_{Flow} (*per hour*)	$AHI_{Chicago\text{-}Cannula}$ (*per hour*)	RDI_{Desat} (*per hour*)	$RDI_{Arousal}$ (*per hour*)	RDI_{Total} (*per hour*)
A. Development set (n = 34)												
No snoring, no EDS	13	1/12	23–65	18–35	0–10	226–429	0–2	0–15	1–17	0–14	1–16	2–17
Snoring, EDS	21	5/16	24–67	21–46	4–19	170–458	0–12	1–33	10–48	2–20	8–46	11–71
B. Test set (n = 103)												
No snoring, no EDS	14	9/5	18–58	20–42	1–7	89–454	0–3	0–17	1–20	0–11	1–19	1–21
Snoring, no EDS	21	3/18	14–71	20–51	0–11	122–457	0–45	0–53	0–55	0–49	0–53	2–57
Snoring, EDS	68	17/51	14–69	21–47	2–24	179–463	0–32	1–43	3–53	0–41	3–49	5–59

Definition of abbreviations: $AHI_{Chicago\text{-}Cannula}$ = apnea + flow hypopnea + flow limitation events with desat/arousal index; AHI_{Flow} = apnea + flow hypopnea index; AI = apnea index; BMI = body mass index; EDS = excessive daytime somnolence; ESS = Epworth Sleepiness Scale; $RDI_{Arousal}$ = apnea + flow hypopnea + flow limitation events with arousal index; RDI_{Desat} = apnea + flow hypopnea with desat + flow limitation events with desat index; RDI_{Total} = apnea + flow hypopnea + flow limitation event index; TST = total sleep time.

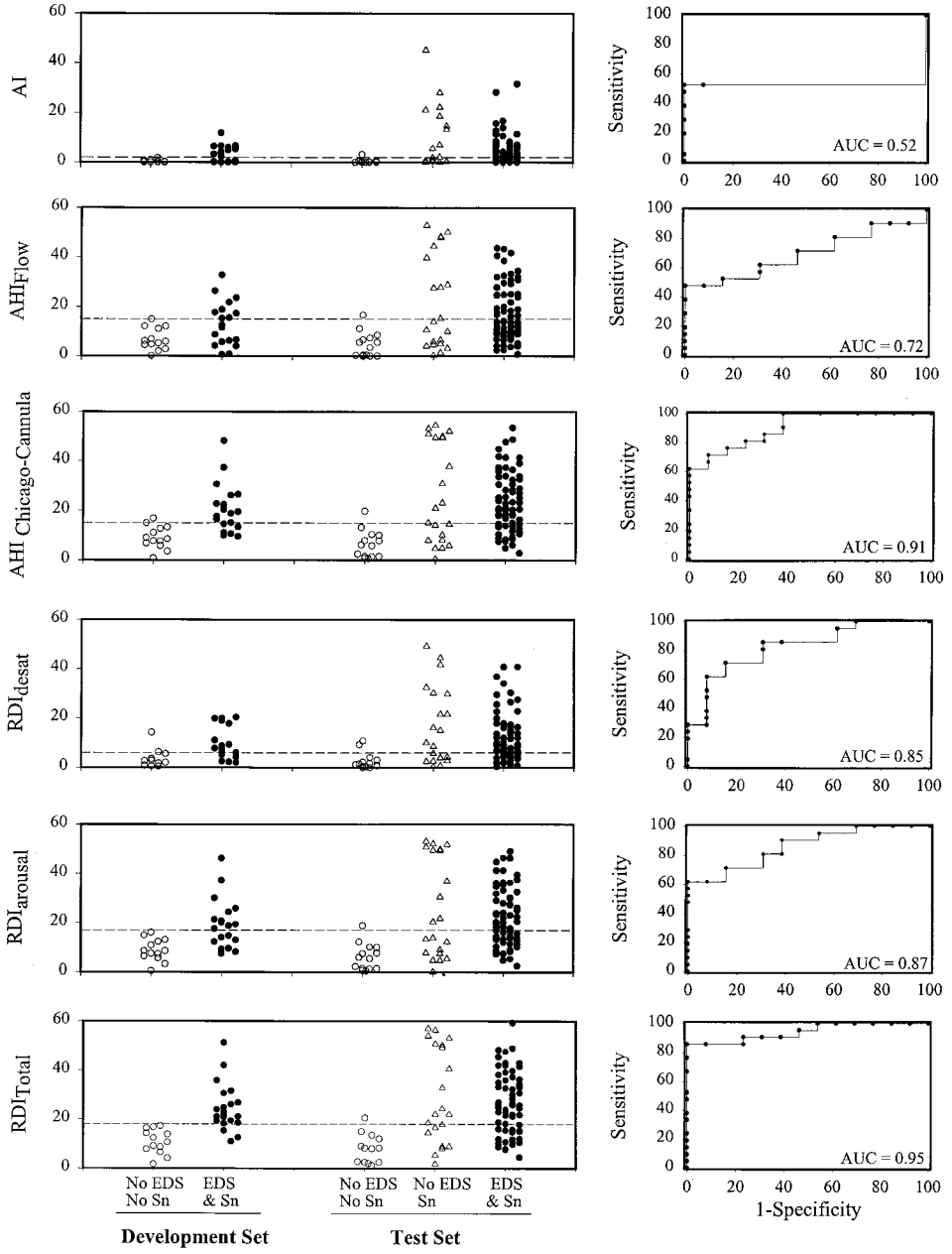

Figure 2. Plot of AI, AHI_{Flow}, $AHI_{Chicago\text{-}Cannula}$, RDI_{Desat}, $RDI_{Arousal}$, and RDI_{Total} for all subjects in the development set (n = 34) and test set (n = 103). Open symbols indicate subjects without excessive daytime somnolence (EDS): nonsnoring (NoSn) subjects (*open circles*), snoring (Sn) subjects (*open triangles*). Subjects with EDS are shown with filled symbols (*filled circles*). On the *right* are the receiver operator curves (ROCs) for each index along with the area under the curve (AUC). The *dashed line* indicates the optimal cut point for each index obtained using the ROC curve.

point. As can be seen from the ROC curves, the RDI_{Total} appeared to have the best discriminant ability (greatest area under the curve, AUC = 0.95). For the ROC curve of RDI_{Total}, the optimal combination of sensitivity and specificity was obtained at a cut point of 18 events/h. In the development set, 18 of 21 subjects with EDS and snoring had an $RDI_{Total} \geq 18$ events/h and 13 of 13 subjects with no EDS/snoring had an $RDI_{Total} < 18$ events/h. This resulted in a sensitivity of 86% and specificity of 100% in the development set for patients thought by clinical assessment (snoring and EDS) to have upper airway dysfunction. It is also apparent from the ROC curves that AI alone and AHI_{Flow} have poor discriminant ability. RDI_{Desat}, $RDI_{Arousal}$, and $RDI_{Chicago\text{-}Cannula}$ have a better discriminant ability, but are still inferior to RDI_{Total}.

Table 3 shows the sensitivity and specificity in the development set for EDS, using the "best" cut point for each index (as identified on the basis of the ROC curves for that index). Also shown are sensitivity and specificity data for other cut points used in common practice. Of note, the cut point or 5/h for the $AHI_{Chicago\text{-}Cannula}$ (proposed by the AASM to be of some relevance) results in a sensitivity of 100% but a specificity for EDS of only 15%.

Using the bidimensional analysis the best cut points (AI > 2 and RDI_{Total}–AI > 16) resulted in a sensitivity of 86% and specificity of 100%, which were identical to those produced by the $RDI_{Total} > 18$ events/h alone. We also examined other statistical classification techniques to separate subjects with EDS from those without EDS (principal components, discriminant analysis, and CART [Systat version 9; SPSS Inc., Chicago, IL] analysis). None of these methods resulted in better separation between groups, and we could not justify their use.

Table 4 shows the sensitivity and specificity for EDS obtained in the test set subjects, using the best indices and cut points obtained from the development set (*see* Table 3). For RDI_{Total}, 48 of 68 subjects with snoring and EDS (symptomatic) fell above the cutoff value of 18 events/h. Twenty-one of

TABLE 3

DEVELOPMENT SET*

Index (*per hour*)	Cut Point	Sensitivity for EDS	Specificity for EDS
AI	2†	52% (11/21)	100% (13/13)
	5	29% (6/21)	100% (13/13)
AHI_{Flow}	5	81% (17/21)	39% (5/13)
	10	57% (12/21)	69% (9/13)
	15†	48% (10/21)	100% (13/13)
$AHI_{Chicago-Cannula}$	5	100% (21/21)	15% (2/13)
	10	90% (19/21)	62% (8/13)
	15†	71% (15/21)	92% (12/13)
RDI_{Desat}	5	81% (17/21)	69% (9/13)
	6†	71% (15/21)	85% (11/13)
	10	38% (8/21)	92% (12/13)
	15	29% (6/21)	100% (13/13)
$RDI_{Arousal}$	5	100% (21/21)	15% (2/13)
	10	81% (17/21)	62% (8/13)
	15	62% (13/21)	92% (12/13)
	17†	62% (13/21)	100% (13/13)
RDI_{Total}	5	100% (21/21)	15% (2/13)
	10	100% (21/21)	46% (6/13)
	15	90% (19/21)	77% (10/13)
	18†	86% (18/21)	100% (13/13)
	20	67% (14/21)	100% (13/13)
AI (RDI_{Total}–AI) bidimensional	2†		
	16†	86% (18/21)	100% (13/13)
	5	86% (18/21)	77% (10/13)
	15		

* n = 34.
† Best.

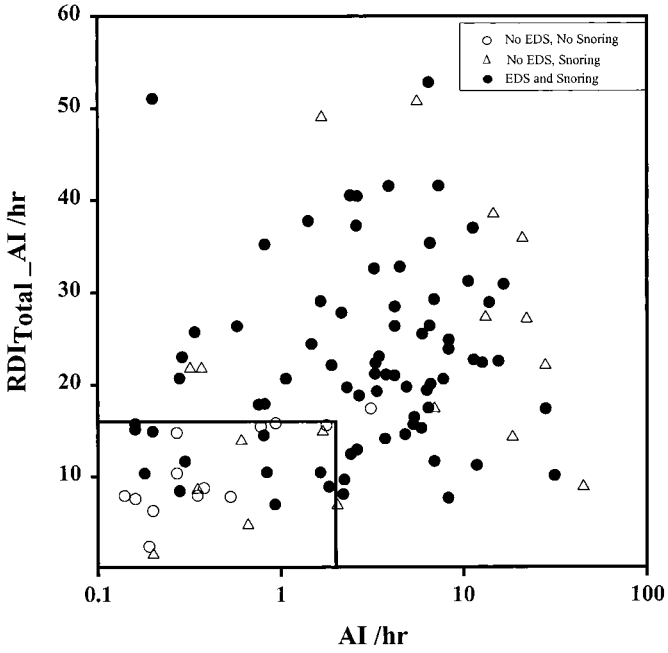

Figure 3. Plot of apnea index (AI) against nonapnea event index (RDI_{Total}–AI) for 103 test subjects. *Open circles* represent asymptomatic subjects, *open triangles* represent snorers without EDS, and *closed circles* represent subjects with EDS and snoring. The region bounded by AI $<$ 2/h and RDI_{Total}–AI $<$ 16/h was defined as identifying the "normal" region.

35 subjects without EDS (asymptomatic) fell below this value. Specifically, 13 of the 14 subjects who had no snoring and no EDS had an $RDI_{Total} < 18$ events/h. The one subject who denied snoring and had no EDS and had an $RDI_{Total} \geq 18$ events/h was observed to have loud snoring during the NPSG. On the other hand, of the 21 snorers who denied EDS, 8 had an $RDI_{Total} < 18$ events/h.

Figure 3 shows the test set data plotted using the two dimensional analysis based on apnea index and RDI_{Total}–AI. The two cut points shown are those derived from the development set.

Shown in Figure 4 is the ESS for all subjects and mean sleep latency (SL) in a subgroup of 32 subjects who had an MSLT. In both the normal group and the snoring subjects who denied EDS, both the ESS and SL are in the "normal" range (ESS $<$ 10 and SL $>$ 10). However, in the group with clinically defined EDS, these indices cover the entire range of values (3–24 for Epworth, 3–20 min for SL) and are frequently not in synchrony. If either test is abnormal, this has predictive value for patient complaints. Despite this, a significant number of individuals will have both tests "normal" and still have clinical complaints of EDS.

TABLE 4

TEST SET*

Index (*per hour*)	Cut Point	Sensitivity for EDS	Specificity for EDS
RDI_{Total}	≥ 18	71% (48/68)	60% (21/35)
	≥ 20	66% (45/68)	63% (22/35)
$AHI_{Chicago-Cannula}$	≥ 5	97% (66/68)	26% (9/35)
	≥ 15	69% (47/68)	63% (22/35)
AI (RDI_{Total}–AI)	≥ 2	79% (54/68)	54% (19/35)
	≥ 16		
AI (RDI_{Total}–AI)	≥ 5	81% (55/69)	51% (18/35)
	≥ 15		

*n = 103.

DISCUSSION

In the upper airway resistance syndrome, as in OSAS, sleep disruption and symptoms of excessive daytime somnolence are thought to result from repetitive respiratory events. Many of these events may be missed by conventional (thermistor-based) monitoring (10, 15). A consensus statement from the AASM has suggested that these otherwise missed events can be detected by esophageal manometry and should be called RERAs. We have previously shown (21) that one can use flow limitation detected on a nasal cannula flow signal to identify these events noninvasively. The present article explores the relationship between various indices of respiratory disturbance and EDS and shows that the simple index based on the total number of apneas, hypopneas, and flow limitation events provides a better sensitivity and specificity than other commonly used indices. This is particularly relevant as the group we analyzed was restricted to patients whose equivocal physiology represents a challenge for the clinician.

The current recommendations (AASM) for defining respiratory abnormality suggest that one should include RERAs in assessing the amount of sleep-disordered breathing, but give little useful guidance as to the cut point between normal and disease if these events are included. Our data indicate that typical cut points of 5 and 10 events/h are too low. This is particularly relevant for the RDI_{Total}, where a cut point of 18 events/h had the best discriminant utility, purely from respiratory information in the NPSG.

Despite optimal separation, 20 of the 68 subjects with EDS in our test set (clinically suspected sleep-disordered breathing) had an RDI_{Total} below the best cut point we could identify

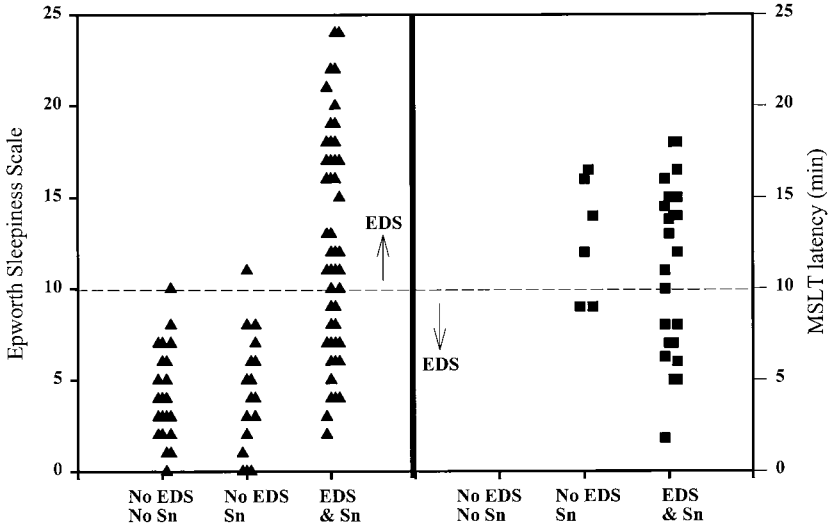

Figure 4. Plot of Epworth Sleepiness Scale (ESS) for all subjects. Sleep latencies for a subset of patients who had an Multiple Sleep Latency Test (MSLT) are shown on the right. The *dashed line* is the conventional cutoff used for EDS: ESS $>$ 10 or MSLT $<$ 10 min.

($<$ 18 events/h). On further review of all available data in these patients, including the NPSG, we found that 9 (45%) of these 20 patients had a coexistent nonrespiratory potential cause for their EDS that was not apparent at the time of initial clinical interview. Examples of these causes were narcolepsy and periodic limb movements (PLMs) not initially in the differential diagnosis. In contrast, the same type of review of the remaining 48 subjects with EDS who had RDI $>$ 18 events/h identified only 8 subjects (17%) who had a coexistent cause of EDS. These data suggest that an RDI_{Total} above the cutoff value of 18 events/h is a strong physiologic correlate of a respiratory basis for the symptom of EDS. Conversely, an RDI_{Total} $<$ 18 events/h has utility in predicting that a subject has a nonrespiratory contribution to his EDS. Thus, our results suggest that the RDI_{Total}, while not perfect, begins to approach the sensitivity required for a screening tool, that is, one whose purpose is to "rule out" significant SDB.

The level of respiratory abnormality (RDI_{Total}) in our 21 snorers without EDS was highly variable. Thirteen of them had RDI_{Total} above our cut point of 18 events/h (range, 18.5–55/h). Not only did these snorers have a total respiratory index similar to that found in individuals who had clear sleep-disordered breathing (who had EDS), but they also had similar numbers of apneas and hypopneas. This suggests that different susceptibility may exist to the same physiologic respiratory stress, at least for EDS and that there is no perfect way to separate individuals completely by counting the respiratory events on the NPSG. Furthermore, we did not find any significant differences in age, sex, weight, or other characteristics in these two groups. Finally, we did not find evidence of the differences between snorers being due to be an error of reporting subjective EDS. Of the 21 snorers without EDS, 8 had either an MSLT or MWT performed and there was no correlation between the RDI_{Total} and the sleep latencies. Only two (RDI_{Total} of 15/h and 54/h) of these eight snorers who reported no EDS and had an MSLT done had sleep latencies $<$ 10 min.

In our analysis, the best classification of our subjects was obtained with a single index, the RDI_{Total}. A similar level of separation was achieved with a bidimensional classification scheme of AI and other nonapnea events (RDI_{Total}–AI). Because of its simplicity, we have emphasized the use of the single RDI_{Total}. Despite this, the bidimensional approach has a conceptual advantage in that different respiratory events may carry different weights in determining symptoms. Thus, a small number of apneas may be more important than a larger number of flow limitation events, and this cannot be reflected in a count of total events. Support for this idea is provided by the observation of Stradling *et al.* that arousals following apneas last longer than those following hypopneas; however, these authors did not relate arousals to symptoms (27). Preliminary data from our laboratory suggest that the conceptually attractive but slightly more intricate bidimensional approach may have advantages in evaluating subjects with positional or sleep stage-dependent SDB (28) or in evaluating change due to partially effective therapy (e.g., dental appliances) (29). This is particularly true when the comparison between two studies involves conversion between types of events (e.g., apnea to hypopnea, flow limitation events or RERAs).

Several additional issues are raised by our methodology. These include (*1*) the definition of EDS and its choice as the outcome measure, (*2*) the effect on sensitivity and specificity of our criteria for selection of subjects and the classification as snorers without EDS as "nondisease," and (*3*) our definition of "mild" respiratory events.

1. Our criterion for the definition of EDS was entirely subjective and relied on patient report and/or the global interpretation of the history by the expert sleep clinician. The MSLT is the acknowledged "gold standard" for assessing sleepiness on the basis of the assumption that physiological need (an increased tendency to fall asleep) leads to symptoms. However, there is a complex relationship between this test and the subjective perception of symptoms that brings a patient to the sleep clinician (30, 31). This is likely to be particularly true in mild disease. The Epworth Sleepiness Scale (ESS) is a self-reported subjective scale used to assess EDS, but this scale has been shown to have only a modest correlation with either MSLT (32) or patient complaints (33), and our own data confirm the lack of a strong relationship. Our goal in the present study was to define an "outcome" measure with which to correlate the physiologic respiratory data from the NPSG. Given the known limitations of the objective tests of physiologic sleepiness and of existing subjective scales, we chose to use as an "outcome" the representation of EDS we believe is most general and perhaps most frequently clinically used, that is, the global assessments by the patient and sleep clinician.

It must be noted that EDS is only one of several outcome variables that could have been used to test the utility of physiologic measurements of SDB. Other studies have raised the

possibility of a link between SDB and hypertension, cardiovascular and cerebrovascular morbidity, and mortality. There is no a priori reason to assume that the same link would exist between an index of SDB and each of these outcomes. In the present study we looked only at EDS, because this variable seems most logically linked to the disruption of sleep experienced by our patient population. If a "best" respiratory index can be identified that correlates with EDS, it would be logical to begin testing this index as an exposure variable for other outcomes, rather than using indices that do not appear to have a tight association with EDS.

2. The selection of our sample is discussed in METHODS. We specifically excluded subjects with EDS and severe obstructive apnea ($AHI_{Flow} > 45/h$) in order to focus our results on the clinically most vexing group (UARS and mild/moderate OSAS). By definition, all subjects eliminated by this restriction would have been true positives in our analysis. The effect of this is that we decreased measured sensitivity in our test set, thus making our results more conservative.

A separate issue affects our analysis of specificity. In our study, specificity was defined for EDS. Of more clinical interest is the specificity of the obtained respiratory indexes and cut points for disease. In this regard, the decision we made to categorize snorers without EDS as nondiseased, as opposed to using snoring as a marker of disease, will affect the calculation of specificity of any respiratory index. At this time, the clinical significance of otherwise asymtomatic snoring is an unsettled issue. It can be argued that a snorer with an AHI > 30 events/h (by any definition of AHI) should be considered as having disease even without the symptom of EDS, rather than being classified as normal because he is asymptomatic. This would convert one of our errors of classification into a correct result and improve specificity of the respiratory index; however, it represents circular reasoning if one wants to test the respiratory metric.

3. The use of the nasal cannula signal and our classification of event types require some discussion. The AASM criteria for sleep-disordered breathing suggest it is appropriate to detect apnea and hypopnea by nasal cannula, and respiratory event-related arousal by esophageal manometry and EEG arousals. Because we have previously shown that an easily identified pattern of changes in shape of the inspiratory airflow derived from the nasal cannula signal (flow limitation event) occurs simultaneously with a classic RERA defined by esophageal manometry, we did not use this technique. Although flow limitation events can occur without arousal (depending on the definition of arousal [20]) flow limitation events terminated by arousal are nearly identical to the standard RERA (21). By one interpretation, the strict application of the AASM rules forces a flow limitation event (defined by a visible change in flow contour most often associated with an amplitude reduction [our unpublished data]) terminated by an arousal to be classified as a hypopnea. However, to accommodate various indices of abnormal respiration and be more consistent with definitions of hypopnea based on other tools, we tabulated flow limitation events separately from hypopnea and defined our summary indices by summing various categories. Only in calculating $AHI_{Chicago\text{-}Cannula}$ were the flow limitation events with arousal reclassified as hypopnea.

Our focus in this data analysis was on defining a metric to quantify sleep-disordered breathing. However, the calculation of an index requires a denominator of time, generally taken as the time of EEG documented sleep. There is a growing interest in systems that propose to detect sleep-disordered breathing without incorporating sleep monitoring (34). We questioned whether it might be possible to use the current data set to evaluate an index calculated without an EEG-based denominator of sleep period. Using the time from lights out to lights on as the denominator for calculating a non-EEG RDI_{Total}, we repeated the ROC analysis in the development set and found that a lower optimal cutoff (15 events/h) resulted in a sensitivity of 86% and specificity of 100%, and a sensitivity of 69% and specificity of 60% in the test set. These results were similar to those obtained with the EEG-based RDI_{Total} with the cutoff of 18 events/h. Thus, in our data set, a non-EEG-based index did not reduce the ability of the best respiratory index to discriminate between disease and normal. However, it is important to note that all of our NPSGs were collected in a monitored setting and that the extrapolation of our conclusions to an unmonitored setting needs further evaluation to determine the effect on the denominator.

In conclusion, our data reemphasize the complex relationship between clinical symptoms and the magnitude of the respiratory disturbance during sleep despite everything done to maximize detection of respiratory events. However, a respiratory index (RDI_{Total}) that includes apneas, hypopneas, and events defined by flow limitation is superior to other indices in its relationship to the clinical symptom of EDS. We propose that this index improves the physiological description of the full spectrum of SDB from normal to UARS and OSAS and may be more useful than the currently defined AHI in evaluating the relationship between SDB and various clinical outcomes.

Acknowledgment: The authors acknowledge the assistance of Harold Dietzius, Rakhil Kanevskaya, Rebecca O'Malley, Alison Rosen, Helen Valentine, and Maureen Wallace in scoring of the records.

References

1. Young T, Palta M, Dempsey J, Skatrud J, Weber S, Badr S. The occurrence of sleep-disordered breathing among middle-aged adults. *N Engl J Med* 1993;328:1230–1235.
2. Guilleminault C, Stoohs R, Clerk A, Cetel M, Maistros P. A cause of excessive daytime sleepiness: the upper airway resistance syndrome. *Chest* 1993;104:781–787.
3. Fleetham JA. A wake up call for sleep disordered breathing. *Br Med J* 1997;314:839–840.
4. Phillipson EA. Sleep apnea—a major public health problem. *N Engl J Med* 1993;328:1271–1273.
5. Guilleminault C, Stoohs R, Shiomi T, Kushida C, Schnittger I. Upper airway resistance syndrome, nocturnal blood pressure monitoring, and borderline hypertension. *Chest* 1996;109:901–908.
6. Silverberg DS, Oksenberg A. Essential hypertension and abnormal upper airway resistance during sleep. *Sleep* 1997;20:794–806.
7. American Thoracic Society. Sleep apnea, sleepiness, and driving risk. *Am J Respir Crit Care Med* 1994;150:1463–1473.
8. Dement WC. The perils of drowsy driving. *N Engl J Med* 1997;337:783–784.
9. Mitler MM, Miller JC, Lipsitz JJ, Walsh JK, Wylie CD. The sleep of long-haul truck drivers. *N Engl J Med* 1997;337:755–761.
10. American Academy of Sleep Medicine. Sleep-related breathing disorders in adults: recommendations for syndrome definition and measurement techniques in clinical research. *Sleep* 1999;22:667–689.
11. Quan SF, Howard BV, Iber C, Kiley JP, Nieto FJ, O'Connor GT, Rapoport DM, Redline S, Robbins J, Samet JM, Wahl PW. The Sleep Heart Health Study: design, rationale, and methods. *Sleep* 1997;20:1077–1085.
12. Kryger MH. Controversies in sleep medicine: terminology and definitions in sleep-disordered breathing. *Sleep* 1997;20:1208.
13. Moser NJ, Phillips BA, Berry DT, Harbison L. What is hypopnea, anyway? *Chest* 1994;105:426–428.
14. Redline S, Sanders M. Hypopnea, a floating metric: implications for prevalence, morbidity estimates, and case finding (review, 26 references). *Sleep* 1997;20:1209–1217.
15. Norman RG, Ahmed MM, Walsleben JA, Rapoport DM. Detection of respiratory events during NPSG: nasal cannula/pressure sensor versus thermistor. *Sleep* 1997;20:1175–1184.
16. Hosselet JJ, Norman RG, Ayappa I, Rapoport DM. Detection of flow limitation with a nasal cannula/pressure transducer system. *Am J Respir Crit Care Med* 1998;157:1461–1467.

17. Redline S, Kapur VK, Sanders MH, Quan SF, Gottlieb DJ, Rapoport DM, Bonekat WH, Smith PL, Kiley JP, Iber C. Effects of varying approaches for identifying respiratory disturbances on sleep apnea assessment. *Am J Respir Crit Care Med* 2000;161:369–374.
18. Loube DI, Andrada T, Howard RS. Accuracy of respiratory inductive plethysmography for the diagnosis of upper airway resistance syndrome. *Chest* 1999;115:1333–1337.
19. Strollo PJJ, Sanders MH. Significance and treatment of nonapneic snoring. *Sleep* 1993;16:403–408.
20. O'Malley EB, Walsleben JA, Norman RG, Rapoport DM. Detection of unappreciated respiratory-related EEG arousals [abstract]. *Am J Respir Crit Care Med* 1996;153:A568.
21. Ayappa I, Norman RG, Krieger AC, Rosen A, O'Malley R, Rapoport DM. 2000. Non-invasive detection of respiratory effort related arousals (RERA**s**) by a nasal cannula pressure/transducer system. *Sleep* 2000; 23:763–771.
22. Johns MW. A new method for measuring daytime sleepiness: the Epworth Sleepiness Scale. *Sleep* 1991;14:540–545.
23. Carskadon MA. Measuring daytime sleepiness. In: Principles and practice of sleep medicine. Kryger MH, Roth T, Dement WC, editors. Philadelphia: W.B. Saunders; 1994. p. 961–966.
24. Rechtschafen A, Kales A. A manual of standardized terminology, techniques, and scoring system for sleep states of human subjects. Washington DC: U.S. Government Printing Office; 1968. NIH Publication No. 204.
25. Guyatt AR, Parker SP, McBride MJ. Measurement of human nasal ventilation using an oxygen cannula as a pitot tube. *Am Rev Respir Dis* 1982;126:434–438.
26. Hosselet JJ, Norman RG, Ayappa I, Oen K, Walsleben JA, Rapoport DM. Use of flow limitation events to characterize sleep disordered breathing [abstract]. *Am J Respir Crit Care Med* 1998;157:A650.
27. Stradling JR, Pitson DJ, Bennett L, Barbour C, Davies RJO. Variation in the arousal pattern after obstructive events in obstructive sleep apnea. *Am J Respir Crit Care Med* 1999;159:130–136.
28. Hosselet JJ, Rapoport DM, Rosen A, Norman RG, Ayappa I. Effect of position and sleep stage on flow limitation events in mild–moderate obstructive sleep apnea syndrome [abstract]. *Sleep* 1998;21:414.K1.
29. Hosselet JJ, Martinez JM, Gelb M, Binder D, Norman RG, Rapoport DM. Utility of identifying flow limitation in assessing the efficacy of oral appliances in mild sleep disordered breathing [abstract]. *Am J Respir Crit Care Med* 1999;159:A248.
30. Chervin RD, Aldrich MS, Pickett R, Guilleminault C. Comparison of the results of the Epworth Sleepiness Scale and the Multiple Sleep Latency Test. *J Psychosom Res* 1997;42:145–155.
31. Chervin RD, Kraemer HC, Guilleminault C. Correlates of sleep latency on the multiple sleep latency test in a clinical population. *Electroencephalogr Clin Neurophysiol* 1995;95:147–153.
32. Benbadis SR, Mascha E, Perry MC, Wolgamuth BR, Smolley LA, Dinner DS. Association between the Epworth Sleepiness Scale and the Multiple Sleep Latency Test in a clinical population. *Ann Intern Med* 1999;130:289–292.
33. Chervin RD, Aldrich MS. The Epworth Sleepiness Scale may not reflect objective measures of sleepiness or sleep apnea. *Neurology* 1999;52:125–131.
34. Stradling JR, Davies RJ. Is it necessary to record sleep? (review, 25 referencess). *Sleep* 1996;19:S251–S254